

AUTOMATISMUS, PLASTIZITÄT UND DIE ABWEICHENDEN URSPRÜNGE DER KÜNSTLICHEN INTELLIGENZ

Biopolitics, GenericScience / 2023-11-30 / Automatismus, Gehirn, Künstliche Intelligenz, Plastizität / Von David Bates

Das heutige Gehirn ist zu einem großen Teil ein digitales Gehirn[1]. Wir erforschen das Gehirn nicht nur mithilfe von Technologien, die auf digitalen Visualisierungen beruhen, sondern auch die Aktivität des Gehirns selbst wird häufig durch digitale Bilder modelliert[2]. Und das Gehirn wird auf unterschiedliche Weise immer noch als eine digitale Maschine, eine Art neuronaler Computer, betrachtet[3]. Das Erbe der Künstlichen Intelligenz (KI) besteht in den zeitgenössischen Neurowissenschaften und Kognitionswissenschaften fort[4]. Die beiden konkurrierenden Projekte zur "Kartierung" des Gehirns in Europa und den USA offenbaren deutlich diskursive und konzeptuelle Verbindungen zwischen Computern, Neurophysiologie und Neuropathologie. Wie auf der Website des Human Brain Project der Europäischen Union zu lesen ist, wird uns das Verständnis des menschlichen Gehirns in die Lage versetzen, "tiefere Einblicke in das zu gewinnen, was uns zu Menschen macht, neue Behandlungsmethoden für Hirnkrankheiten zu entwickeln und bahnbrechende neue Computertechnologien zu bauen"[5]. Auch das vom National Institute of Health initiierte und von Präsident Obama mit erheblichen Mitteln geförderte Connectome-Projekt offenbart den wesentlichen Zusammenhang zwischen der Erforschung neuer Technologien, der technologisierten Sicht auf das Gehirn selbst und den Pathologien : "Lange ersehnt von Forschern, die nach neuen Wegen suchen, um Hirnfunktionsstörungen zu behandeln, zu heilen und sogar zu verhindern, wird dieses Bild wichtige Lücken in unserem derzeitigen Wissen schließen und beispiellose Möglichkeiten bieten, genau zu erforschen, wie das Gehirn es dem menschlichen Körper ermöglicht, große Mengen an Informationen aufzunehmen, zu verarbeiten, zu nutzen, zu speichern und abzurufen – und das alles mit der Geschwindigkeit des Denkens."[6] Das Gehirn

ist der Ort, an dem der menschliche Körper die Informationen aufnimmt, verarbeitet, nutzt, speichert und abruft.

Es scheint, dass die jüngste Intensivierung des Interesses an der inhärenten Plastizität des Gehirns – seine Entwicklungsbereitschaft, seine sich in der Erwachsenenphase ständig verändernde Struktur und seine erstaunliche Fähigkeit, sich nach einem schweren Trauma neu zu organisieren – eine erhebliche Herausforderung für die technologischen Konzeptualisierungen des Gehirns darstellt, die von einer komplexen, aber definitiv automatischen Funktionsweise ausgehen. In der Tat wurde das Konzept der Plastizität insbesondere von der Philosophin Catherine Malabou[7] als Hindernis für ein maschinelles Verständnis des Gehirns dargestellt. Heute stellt sich jedoch heraus, dass das neurophysiologische Phänomen der Gehirnplastizität schnell mit Computermodellen und digitalen Darstellungen des Nervensystems gleichgesetzt wurde. Beispielsweise wird das Gehirn im Forschungsbereich der Computational Neuroscience als lernende Maschine (learning machine) dargestellt, die nach vordefinierten Algorithmen automatisch Verbindungsnetze aufbaut, die von der spezifischen “Erfahrung” des Netzes abhängen. Diese Modelle leiten sich alle von der grundlegenden Theorie des Neurophysiologen Donald Hebb ab. Wie er in seinem Buch The Organization of Behavior aus dem Jahr 1949 erklärte, werden synaptische Verbindungen zwischen Neuronen durch Gebrauch verstärkt; diese Theorie wird oft in der Formel “Neuronen, die sich gemeinsam erregen, verknüpfen sich miteinander” zusammengefasst. Die Prozesse, die die Konfiguration des Plasthirns im Zuge seiner Erfahrungen mit der Welt steuern, sind natürlich viel komplexer, aber das Grundprinzip bleibt gültig. Deshalb kann sogar die kontingente Konfiguration des Plastic Brain, so glaubt man, durch eine Computersimulation rigoros modelliert werden.

Diese Verdoppelung des Gehirns durch sein digitales Gegenstück hat sich wiederum auf den Bereich der Informatik ausgewirkt. Versuche, die Hirnrinde von Tiergehirnen mit “synaptischen” Architekturen zu modellieren, werden beispielsweise als experimentelle Untersuchungen der neuronalen Organisation selbst dargestellt; es heißt, dies sei ein “wichtiger Schritt zur Aufklärung des Geheimnisses, was das Gehirn berechnet”, was wiederum, so die Forscher, “den Weg zu den kompakten, stromsparenden neuromorphen und synaptronischen Computern von morgen ebnet”[8]. Die digitalen Visionen des plastischen Gehirns haben die Erfindung neuer Computerarchitekturen angeregt. IBM hat zum Beispiel einen Chip entwickelt, der die Netzwerkkonnektivität neuronaler Systeme nachahmt. Dieser Chip ist mit “digitalen Neuronen” ausgestattet, die Synapsen entsprechend ihrer Reaktionen (Inputs) neu verbinden können[9]. Ein Konkurrenzprojekt von Intel verspricht ebenfalls programmierbare “Synapsen” und die Nachahmung von “integrierenden Neuronen”[10]. Diese neurosynaptischen Chips sind Teil einer umfassenderen Untersuchung des sogenannten “Cognitive Computing”, dessen Ziel es ist, einen kohärenten, einheitlichen und universellen Mechanismus zu entwickeln, der von den Fähigkeiten des Geistes inspiriert ist. Mit anderen Worten: Die Forscher “streben nichts Geringeres an, als die grundlegenden Algorithmen des Gehirns zu entdecken, zu demonstrieren und bereitzustellen und ein tiefes wissenschaftliches Verständnis davon zu erlangen, wie der Geist wahrnimmt, denkt und handelt”. Hier ist das Gehirn zu einer algorithmisch lernenden Maschine (algorithmic learning machine) geworden; es ist der Spiegel der Technologie selbst, die es repräsentiert. Die Analyse der vernetzten Organisation des Gehirns wird, so heißt es, “zu neuen kognitiven Systemen, neuen Computerarchitekturen, neuen Programmierparadigmen, neuen praktischen Anwendungen und sogar zu intelligenten Wirtschaftsmaschinen führen”[11].

Dieser Spiegel hat in der Tat eine lange Geschichte, die wahrscheinlich genauso lang ist wie die Geschichte der Informatik selbst. Die frühe KI-Forschung verfolgte zwei Ziele: eine intelligente Simulation zu produzieren und dabei Hypothesen über die Funktionen des menschlichen Geistes zu testen. Dieses Vorhaben beruhte auf der Annahme, dass sich der Geist oder das Gehirn für eine solche Analyse eignen würde. Wie die Forscher 1954 über ihre eigenen Bemühungen, die Großhirnrinde zu modellieren, schrieben: “Wie komplex auch immer ein System organisiert ist […], es kann immer so genau wie gewünscht von einem digitalen Computer simuliert werden, vorausgesetzt, seine Organisationsregeln sind bekannt.”[12] Die Forscher waren sich einig, dass die Großhirnrinde eine wichtige Rolle in der Entwicklung des Gehirns spielte. Dies war, wie sie anmerken, die Implikation von Alan Turings ursprünglicher Konzeptualisierung des universellen digitalen Computers als einer Maschine, die (virtuell) jede andere Maschine im diskreten Zustand imitieren kann.

Die besagte Automatik der Gehirnfunktion (selbst in ihrer radikalsten plastischen Form) ist zum großen Teil eine Folge der historischen Ko-Entwicklung von Computertechnologie, KI und Kognitionswissenschaften. Die Frage, was das Konzept der Automatizität in nur einem dieser Bereiche unter Ausschluss der anderen sein könnte, ist hier bedeutungslos. Sobald also die Plastizität vollständig in die Computer- und neurophysiologischen Modelle des Gehirns integriert ist, kann sich der Widerstand gegen die vollständige Automatisierung des menschlichen Denkens nicht einfach auf romantisierende Konzepte der persönlichen Identität oder auf philosophische Kritik am Materialismus stützen; vielmehr müssen wir uns auf die historischen und konzeptuellen Grundlagen des digitalen Gehirns selbst konzentrieren. Kritik kann durch eine kritische Geschichte des Automatismus entstehen. Wenn wir an die Schwelle des digitalen Zeitalters zurückkehren, zu dem Zeitpunkt, als der moderne Computer zum ersten Mal konzeptualisiert wurde (und gleichzeitig die Ideen und Praktiken der KI in

Bewegung gesetzt wurden), können wir sehen, dass das Digitale ursprünglich nicht vollständig auf die Automatizität ausgerichtet war. Denn obwohl dem nicht viel Aufmerksamkeit geschenkt wurde, versuchten die Hauptentwickler der frühen Computertechnologien ausdrücklich, mit ihren neuen Technologien eine radikalere Offenheit, eine unvorhersehbarere Plastizität innerhalb des Nervensystems nachzuahmen – ein Thema, das, wie wir sehen werden, in der Neurophysiologie des frühen zwanzigsten Jahrhunderts und in der neurologisch begründeten experimentellen Psychologie sehr lebendig war. Zu einer Zeit, als einige Kybernetiker behaupteten, das Gehirn sei nur ein automatischer Rechner wie der Computer, erkannten wichtige Persönlichkeiten in der Geschichte der Informatik und Kybernetik sofort die Bedeutung der Plastizität des Gehirns für das KI-Projekt: Das plastische Gehirn, so glaubte man, biete die Möglichkeit, kreative und unvorhersehbare Sprünge der menschlichen Intelligenz zu modellieren, Fähigkeiten, die über die unerbittlich automatische Ausführung starrer Funktionsmechanismen oder gewohnter Verhaltensweisen hinausgehen. Es ist daher bezeichnend, dass die damaligen neurophysiologischen Diskurse über Plastizität eng mit den Störungen und Krisen des kranken oder verletzten Gehirns verbunden waren. Die Konstruktion einer plastischen Computermaschine zu Beginn des digitalen Zeitalters implizierte daher meiner Meinung nach die Erfindung einer maschinellen Pathologie.

Die Rekonstruktion dieser Geschichte wird eine neue Perspektive auf unser zeitgenössisches "digitales Gehirn" eröffnen. Wir dürfen nicht auf bloße lernende Maschinen reduziert werden, die sich ihrer eigenen kognitiven Prozesse weitgehend unbewusst sind und in denen jede Erfahrung von Freiheit und Kontinuität des Denkens als eine Art nietzscheanischer Illusion entlarvt werden kann. Das menschliche Gehirn wurde als eine besondere Art von wirklich offenem System betrachtet, das sich selbst bestimmt, aber auch in der Lage ist, seinen eigenen Automatismus in Akten echter Kreativität herauszufordern. Die Schöpfer

des digitalen Computers haben sich bei ihren Bemühungen, die Fähigkeiten wirklich intelligenter Wesen zu modellieren, ausdrücklich an diesem neurophysiologischen Konzept der Plastizität orientiert.

Die Geschichte der KI geht zumeist auf Denker zurück, die versuchten, die Fähigkeit zu denken durch eine Art Automaten zu simulieren. Man könnte sagen, dass die philosophischen und physiologischen Schriften von René Descartes uns eine Vision des ersten modernen Automaten, d. h. einer Denkmaschine, vermittelt haben. Obwohl er sich offensichtlich den ultimativen Implikationen seiner eigenen systematischen Mechanisierung des menschlichen und tierischen Körpers widersetzte, ebnete Descartes den Weg für ein mechanisches Verständnis der Kognition, als er eine umfassende Beschreibung des Nervensystems und der entscheidenden Rolle lieferte, die es nicht nur bei der Steuerung sämtlicher tierischer Verhaltensweisen spielte, sondern – was noch wichtiger ist – auch den weitaus größten Teil des menschlichen Denkens und Handelns hervorbrachte, die routinemäßige Kognition des Alltagslebens.

Doch wenn man sich Descartes' Schriften genauer ansieht, entdeckt man eine Theorie, die etwas komplexer ist als eine einfache mechanische Sicht auf den Körper. Descartes zeigte, wie die Nerven und das Gehirn ein Informationssystem bilden, das durch seine Flexibilität und Anpassungsfähigkeit bemerkenswert ist. Auch wenn das Nervensystem eine materielle Struktur ist, ist es dennoch plastisch und veränderbar. Als Raum für die Integration von Informationen war das Gehirn, wie er schrieb, von "weicher und faltbarer"[13] Zusammensetzung und daher in der Lage, von Erinnerungen durchdrungen zu werden und Reflexe zu erwerben. Die Implikation war, dass die kulturelle Bestimmung des Einzelnen – durch Sprache, Kultur und Geschichte – innerhalb der flexiblen Architektur des Gehirns selbst stattfand. Wie er im Discours de la méthode bemerkte, "habe ich auch gedacht, dass derselbe Mensch mit demselben Geist, wenn er von Kindheit

an unter Franzosen oder Deutschen erzogen wird, sich anders entwickelt, als wenn er immer unter Chinesen oder Kannibalen gelebt hätte"[14]. Das offene Gehirn wurde von den physischen Strömen bestimmt, die von den Sinnesorganen erzeugt und durch die Nervenleitungen weitergeleitet wurden. Descartes' Automat war kein Uhrwerk, sondern ein offener Ort der ständigen Organisation und Reorganisation, die durch die von den Sinnessystemen empfangenen Informationen hervorgerufen wurde[15].

Es stimmt, dass die lange Geschichte der Technologien des automatisierten Denkens zeigt, dass die Idee eines "weichen und flexiblen" Systems von den starren Mechanismen des Industriezeitalters in den Schatten gestellt wurde. Von den halbautomatischen Rechenmaschinen von Pascal und Leibniz über Charles Babbages Difference Engine (das erste Beispiel für eine wirklich automatische Form des künstlichen Denkens) und den unvollendeten Generalcomputer, den er Analytic Engine nannte, bis hin zu den "logischen Klavieren" des späten 19. Jahrhunderts (die beispielsweise von William Jevons und Allan Marquand konstruiert wurden) wurden die Modelle des menschlichen Denkens konkret durch Maschinen mit präzisen und vorhersehbaren Operationen verkörpert. Es ist daher nicht überraschend, dass diese frühen Beispiele für künstlich mechanisiertes Denken die Konzeptualisierung der menschlichen Kognition stark beeinflusst haben. So stellte beispielsweise der amerikanische pragmatistische Philosoph Charles Sanders Peirce, der über logische Maschinen nachdachte, fest, dass die Ausführung einer deduktiven Inferenz "keine Beziehung zu dem Umstand hat, dass die Maschine mit Zahnrädern arbeitet, während der Mensch mit einer missverstandenen Anordnung von Gehirnzellen arbeitet"[16]. Peirce selbst hatte (viele Jahre vor Claude Shannon) die Idee, dass diese logischen Beziehungen sogar mithilfe von elektrischen Schaltkreisen realisiert werden könnten[17]. Mit diesem Schritt befinden wir uns an der Spitze der modernen Informatik und

damit auch des Bemühens, das Gehirn selbst als Denkmaschine zu verstehen, die aus einer Vielzahl von neuronalen "Schaltern" aufgebaut ist.

Wir müssen jedoch eine Pause einlegen und anmerken, dass vor dem Aufkommen der digitalen Informatik das Wissen über das Gehirn keineswegs dem mechanistischen Paradigma des Automatismus entsprach. Jahrhunderts wich die Vorstellung, das Gehirn bestehe aus einer Vielzahl lokalisierbarer Komponenten, Modellen des Gehirns, die seine verteilte strukturelle Organisation betonten. Der große britische Neurologe John Hughlings Jackson beispielsweise stützte sich auf umfassende klinische Studien neuronaler Störungen, um zu argumentieren, dass das Gehirn aufgrund der fortschreitenden Entwicklung des Organs im Laufe der Evolution mehrere Organisationsebenen aufweist. Interessant ist, dass Hughlings Jackson nachweisen wollte, dass die höchsten intellektuellen Fähigkeiten mit den am wenigsten organisierten, am wenigsten "entwickelten" kortikalen Strukturen einhergehen, die er als "am wenigsten automatisch" bezeichnete. Die eher automatischen Funktionen regeln unsere grundlegenden physiologischen Systeme, während der relativ offene und unbestimmte Kortex mit komplexem Denken und diskursiver Sprache in Verbindung gebracht wird.

William James, die vielleicht einflussreichste Figur in der Synthese von Psychologie und Neurowissenschaften um die Jahrhundertwende, interessierte sich sehr dafür, wie der menschliche Geist von gewohnheitsmäßigen und unbewussten Kognitionen gesteuert wurde. Diese waren nicht gegeben, sondern vielmehr erlernt. Das Gehirn war sowohl ein Ort der Offenheit als auch ein Raum für künstliche Mechanismen: "Die Phänomene der Gewohnheit bei Lebewesen sind auf die Plastizität der organischen Materialien zurückzuführen, aus denen ihre Körper bestehen". Plastizität", schreibt James, "ist der Besitz einer Struktur, die schwach genug ist, um einem Einfluss nachzugeben, aber stark genug, um nicht auf einmal nachzugeben". Für James wie auch für Peirce sind die

organisatorischen Formen des Lebens im Wesentlichen plastisch, da sie das Produkt evolutionärer Veränderungen sind, und das Nervensystem ist der bevorzugte Ort dieser Neu- und Umbildung, während sich der Organismus an seine wechselnden Umweltbedingungen anpasst. Er stellte fest: “Organische Materie, insbesondere Nervengewebe, scheint mit einem außerordentlichen Maß an Plastizität ausgestattet zu sein”, und erklärte damit, wie Wesen zu einer Art Automaten werden können, aber nie zu echten Automaten, eben weil die Plastizität nie ganz aus dem Gehirn und dem Nervensystem verschwindet[18].

In Anlehnung an evolutionäre Theorien wie die von Herbert Spencer und Hughlings Jackson greift James die Idee auf, dass die menschliche Fähigkeit zu Intuition und Scharfsinn höchstwahrscheinlich nicht mit einem bestimmten Organ der Intelligenz, sondern tatsächlich mit der Instabilität und Unbestimmtheit der oberen Gehirnzentren zusammenhängt. James schlägt vor, dass “gerade die Unbestimmtheit [der Gehirnhälften] ihren Vorteil ausmacht [...]. Ein Organ, das von leichten Eindrücken beeinflusst wird, ist ein Organ, dessen natürlicher Zustand ein instabiles Gleichgewicht ist”. Diese “Plastizität” des höheren Gehirns macht es unberechenbar; es ist “zu einem bestimmten Zeitpunkt genauso anfällig für Wahnsinn wie für Vernunft”. James ging jedoch noch weiter und stellte fest, dass ein verletztes Gehirn, dem Teile fehlen, die beschädigt oder völlig zerstört wurden, als “eine virtuell neue Maschine” betrachtet werden kann. Diese neue Maschine kann anfangs abnormal funktionieren und psychopathologische Symptome aufweisen, aber wie neurologische Fälle immer wieder gezeigt haben, findet dieses neue Gehirn oft den Weg zurück in die Normalität. “Ein Gehirn, dem ein Teil entfernt wurde, ist praktisch eine neue Maschine, und in den ersten Tagen nach der Operation funktioniert es völlig anormal. Allerdings wird seine Leistung von Tag zu Tag normaler, bis es eines geschulten Auges bedarf, um etwas Abnormales zu vermuten.”[19] James war der Meinung, dass die Gehirnforschung in der Lage sein müsse, das Problem zu lösen. Auch James’ Kollege Peirce hatte

einen Zusammenhang zwischen Pathologie und Neuheit hergestellt. Wenn unser Gehirn geschädigt ist", so Peirce, "handeln wir auf eine neue Art und Weise". Normalerweise führt dies zu einem Zustand der Geisteskrankheit, aber Peirce war bereit zuzugeben, dass in besonderen Fällen "eine Erkrankung des Gehirns zu einer Verbesserung der allgemeinen Intelligenz führen kann"[20]. Die dem Gehirn innewohnende Plastizität zeigt sich vor allem bei Hirnverletzungen, aber die Fähigkeit, sich neu zu organisieren und sich gegen Schäden zu verteidigen, ist mit der kognitiven Flexibilität des menschlichen Geistes verbunden. Intelligenz wurde in gewissem Sinne als Folge einer gewissen Desorganisation und Unberechenbarkeit angesehen, und selbst eine pathologische Desorganisation konnte die Leistungen einer genialen Intelligenz erklären.

Zu Beginn des zwanzigsten Jahrhunderts erforschte die klinische und experimentelle Hirnforschung systematisch die Fähigkeit des Gehirns, sich angesichts von Herausforderungen, einschließlich der radikalen Herausforderung einer schweren Verletzung, neu zu organisieren. Die Entdeckung der Unordnung eröffnete die Möglichkeit einer neuen Form der Organisation, die in rein mechanischen Begriffen nicht erklärbar war. Wie Constantin von Monakow in einem Buch über Gehirnkartografie sagt, führt die Störung eines Teils des Gehirns zu einem allgemeineren "Schock" des Systems (den er Diaschisis nennt). "Jede Schädigung der Hirnsubstanz wird (wie die Schädigung jedes anderen Organs) zu einem Kampf um die Erhaltung der gestörten Nervenfunktion führen, und das Zentralnervensystem ist immer (wenn auch nicht immer in gleichem Maße) auf einen solchen Kampf vorbereitet"[21]. Die pathologische Abweichung ruft eine totale Reaktion hervor, die nicht auf eine einfache Rückkehr zur ursprünglichen Ordnung abzielt, sondern vielmehr auf eine Ordnung, die eine stabile Funktion in einer neuen Form wiederherstellt. Wie von Monakow (in einem mit dem französischen Neurologen R. Mourgue verfassten Buch) schreibt, "handelt es sich um einen Kampf, ein aktives Ringen um die Schaffung eines neuen Zustands, der

eine neue Anpassung des Individuums und seiner Umwelt ermöglicht. Der gesamte Organismus greift in diesen Kampf ein”[22].

Die neurologische Theorie der Zwischenkriegszeit brachte zahlreiche theoretische Modelle hervor, die die Plastizität des Nervensystems betonten, das in der Lage ist, sich an ständig wechselnde Umstände und sogar an radikale Hirnschäden anzupassen[23]. Eine einflussreiche Persönlichkeit auf diesem Gebiet, der amerikanische Psychologe Karl Lashley, versuchte, die Theorien über die Lokalisierung des Gedächtnisses zu widerlegen, indem er systematisch Teile des Gehirns von Tierversuchen entfernte und die Persistenz gelernter Verhaltensweisen in Labyrinthen und anderen Umgebungen nachwies. Überraschenderweise waren die getesteten Tiere oft noch in der Lage, die Labyrinthe effizient zu durchlaufen, was impliziert, dass das Gehirn einen Weg haben muss, sich umzustrukturieren, um das fehlende Gewebe zu kompensieren. Für Lashley ist diese Fähigkeit eine Fortsetzung der adaptiven Einheit des Gehirns als komplexes System, ein System, das Aktivität in alle seine Komponenten integriert: “Lashleys Begriff für diese Flexibilität war Äquipotentialität, in Anlehnung an seine Theorie, dass das Gehirn als dynamische, sich ständig verändernde Einheit sich neu organisieren kann, wenn es mit neuen Hindernissen oder internen Fehlern konfrontiert wird.”[24] Lashley war der Ansicht, dass das Gehirn in der Lage ist, sich selbst zu reorganisieren, wenn es mit neuen Hindernissen oder internen Fehlern konfrontiert wird, und dass es in der Lage ist, sich selbst zu reorganisieren, wenn es mit neuen Hindernissen konfrontiert wird. Wie er später schrieb, “war ich beeindruckt von der Tatsache, dass selbst eine sehr starke Zerstörung des Gehirngewebes keine Desorganisation hervorruft. Das Verhalten vereinfacht sich, bleibt aber adaptiv”. Für Lashley hing die intuitive, scharfsinnige und intellektuelle Aktivität von dieser Eigenschaft des Gehirns und des Nervensystems ab, nämlich “seiner plastischen und adaptiven Natur”[25].

Die vielleicht wichtigste Figur in dieser Geschichte der Plastizität war Kurt Goldstein, der seine eigenen umfangreichen klinischen Forschungen zu neurologischen Defekten bei Patienten mit Hirnverletzungen mit breiteren Denkströmungen über das Organisationsleben selbst zusammenführte. Gestützt auf neurologische Konzepte wie die Diaschisis aus der Arbeit von von Monakow sowie auf die deutsche holistische Philosophie, insbesondere die Gestalttheorie, definierte Goldstein den Organismus weniger als eine feste Organisation oder Struktur, sondern vielmehr als eine dynamische Konfiguration, die ständig darum kämpft, angesichts der sich ständig ändernden Lebensbedingungen eine kohärente Einheit zu bewahren. Pathologische Zustände des Wesens weisen bei Patienten, die neurologische Störungen zeigen, ihre eigenen Merkmale auf. Wenn die Reaktionen eines Organismus auf die Umwelt "konstant, korrekt, angemessen" sind, kann er in seiner Umgebung überleben. In einem Schockzustand wird der Organismus jedoch häufig zu "ungeordneten, unbeständigen, inkohärenten" Handlungen veranlasst, was zu einer "beunruhigenden Folge"[26] führt.

Wie Goldstein behauptet, haben Organismen jedoch die Fähigkeit, sich angesichts dieses unkoordinierten Verhaltens zu verändern, was er als "Katastrophenreaktion" bezeichnet. Die Fähigkeit des Organismus, sich als Reaktion auf einen Schock neu zu organisieren, ist laut Goldstein nur ein Mittel, mit dem der Organismus in Zeiten der Störung nach Einheit strebt. Die als pathologisch bezeichneten Katastrophen sind in Wirklichkeit Teil der Kontinuität der Dynamik eines gesunden und normalen Organismus, der sich bemüht, sein Gleichgewicht zu erhalten. Mit anderen Worten, der Organismus ist immer auf der Suche nach Einheit und Ordnung: "Normale und abnormale Reaktionen ('Symptome') sind nur der Ausdruck des Versuchs des Organismus, mit bestimmten Anforderungen der Umwelt fertig zu werden"[27].

So schreibt Goldstein, dass selbst das normale und gesunde Leben des Organismus als eine Reihe von – wie er es nennt – “leichten Katastrophenreaktionen” betrachtet werden kann, bei denen man zunächst mit Unzulänglichkeiten konfrontiert wird und dann nach “neuen Anpassungen” oder einem “neuen geeigneten Milieu” sucht, um diesem Mangel zu begegnen. Der katastrophale Bruch, der schwerwiegender ist, unterscheidet sich nur durch das Ausmaß und die Intensität der Reaktion. Der gesamte Organismus ist in Wirklichkeit eine Einheit in spe, die immer wieder in Schockzustände gerät, und er muss immer und immer wieder auf kreative Weise eine neue Ordnung herstellen, um diese Schocks zu überwinden: “Wenn der Organismus “sein” soll, muss er immer wieder von Momenten der Katastrophe zu Zuständen geordneten Verhaltens übergehen.”[28] Der Organismus ist eine Einheit in spe, die immer wieder in Schockzustände gerät, und er muss immer wieder auf kreative Weise eine neue Ordnung herstellen, um diese Schockzustände zu überwinden. Die wesentliche Plastizität des Organismus ist im Gehirn natürlich am deutlichsten sichtbar, weil es das Zentrum der Organisation und Integration und damit der Ort der Neuorganisation und Wiederherstellung ist.

Die gleichen Strukturen von Krise und Reorganisation lassen sich jedoch auch im Bereich der Psychologie der Zwischenkriegszeit ausmachen, insbesondere in den Theorien des kreativen und adaptiven Denkens. In der Gestaltpsychologie wurde wahre Intelligenz als die Fähigkeit definiert, schwierige Probleme mit neuen Perspektiven zu lösen, die die Erfahrung des Subjekts reorganisieren, um neue Arrangements hervorzubringen. Das Wort, das für diese plötzliche Erkenntnis verwendet wird, ist insightb. In seinen berühmten Affenstudien während des Ersten Weltkriegs versuchte Wolfgang Köhler, die Urbedingungen der Intelligenz zu entdecken, indem er Primaten als Mittel benutzte, um zum Wesen der menschlichen Intelligenz vorzudringen, bevor sie von Sprache und kulturellem Wissen überlagert wurde.

Köhler zeigte, dass Einsicht (insight) erreicht wird, wenn der Geist sich von einer Interpretation der Situation befreit, um eine neue zu entdecken, die die Spannung auflöst – in diesen Fällen Gegenstände als Werkzeuge zu sehen, die zur Nahrung führen können, die außerhalb des direkten Zugriffs platziert worden war. Für Köhler offenbart sich Intelligenz durch die Fähigkeit, vom direkten Weg "abzuweichen", um zu verstehen, wie Hindernisse umgangen werden können[29]. Die psychologische Einsicht hat ihr "isomorphes" Gegenstück im Organismus selbst: "Menschen bauen Maschinen und zwingen die Naturkräfte, in einer bestimmten, kontrollierbaren Weise zu funktionieren. Ein solches System besteht aus "festen" Verbindungen im Sinne der analytischen Mechanik". Ein dynamisches System ist jedoch in der Lage, sich neu zu organisieren und sich anzupassen. Köhlers Beispiel ist der sich entwickelnde Organismus: "Viele Embryonen können vorübergehend (künstlich) auf eine Art und Weise deformiert werden, die ihrer genetischen Geschichte völlig unangemessen ist, und dennoch auf einem völlig anderen Weg zu demselben Entwicklungsstand zurückkehren, den ein anderer, intakter Embryo erreicht hat.

Die Untersuchung des "produktiven Denkens" betonte die Neuartigkeit der Analogieprojektionen von zuvor etabliertem Wissen in radikal neue Verständnisbereiche[31]. Köhlers Kollege und ebenfalls Anhänger der Gestalttheorie, Kurt Koffka, argumentierte zum Beispiel, dass der Geist nicht auf die Kombination und Neukombination von "Reaktionswegen" und Erinnerungen, die er bereits besaß, beschränkt sei. Der Geist besitzt seiner Meinung nach eine "grundlegende Plastizität", die radikale Neuheit und damit einen echten Fortschritt im Denken ermöglicht[32]. Der Schweizer Experimentalpsychologe Edouard Claparède hat diese Beziehung zwischen den Automatismen des gewohnheitsmäßigen Denkens und der für die Intelligenz charakteristischen radikalen Freiheit perfekt erfasst; er knüpfte auch an Goldsteins Idee der katastrophalen Reaktion an: "Die Störung des Gleichgewichts, wenn Reflexe oder

Gewohnheiten nicht zum Eingreifen zur Verfügung stehen, wird nicht automatisch wiederhergestellt, und wir sind vorübergehend nicht angepasst. Es ist die Intelligenz, die dafür sorgt, dass wir uns wieder anpassen”[33].

An der Schwelle zum Computerzeitalter ging man davon aus, dass die offene Struktur des Gehirns eine einfache Reduzierung der menschlichen Kognition auf bestimmte Formen einer automatischen “Maschinerie” ausschließt. Dennoch waren die fortschrittlichsten Computergeräte der damaligen Zeit analoge Maschinen, wie der Differentialanalysator von Vannevar Bush (ca. 1930) und der elektromechanische Hybrid Mark I von Harvard (1940-44). So komplex diese technischen Maschinen auch waren, verkörperten sie doch eindeutig das Prinzip der Automatik: Aufgebaut als eine Reihe physisch gekoppelter beweglicher Elemente, waren diese analogen Geräte darauf ausgelegt, die Interaktionen verschiedener mathematischer oder anderer bestimmter Funktionen direkt darzustellen. Dies schien jede ernsthafte Form von KI auszuschließen, die intelligentes Leben nachahmen kann. Wie Vladimir Jankélévitch 1931 in einem Buch über Bergson sagte: “Bei einer perfekten Maschine gibt es nie eine Desillusionierung, aber auch nie eine Überraschung. Keines dieser Wunder, die in gewisser Weise die Signatur des Lebens sind”[34].

Als Turing sich also 1936 in einem mathematischen Artikel zum ersten Mal den “universellen” Digitalcomputer vorstellte, handelte es sich um einen seltsamen Eindringling in das Gebiet des maschinellen Rechnens, eben weil er keine eigentliche Struktur besaß. Als einfache Maschine, die nicht mehr als die Fähigkeit besaß, zwei Symbole auf einem sich bewegenden Band zu manipulieren, wurde dieser digitale Computer als eine radikal offene Maschine definiert, deren einziger Zweck darin bestand, die Konfigurationen anderer zustandsdiskreter Maschinen zu übernehmen. Was Turing erfunden hatte, war eine Maschine, die automatisch (virtuell, in verschlüsselten Darstellungen) die aufeinanderfolgenden

Zustände einer anderen automatischen Maschine nachahmt. Sehr schnell wurde aus Turings Maschine natürlich viel mehr als das. Die binäre Logik des digitalen Computers wurde bald auch auf die synaptische Konnektivität im Gehirn angewendet. Turings Hauptargument, dass die Funktionsweise eines digitalen Computers von seiner logischen und nicht von seiner physischen Organisation bestimmt wird, hat die Analogie zwischen Computer und Gehirn nur noch weiter verstärkt. Für einige war das Gehirn seinem Wesen nach ein digitaler Computer, der durch die sequenzielle Reaktion von Neuronen instanziiert wurde, die analog zum mechanischen Anschalten von Relais oder der Verarbeitung elektrischer Impulse waren. Das physikalische Substrat der Turingmaschine war für ihre logische Nutzung irrelevant[35].

Die Kritiker der mechanistischen Weltsicht konzentrierten sich auf die starre Automatik, die von der Computeranalogie impliziert wurde. “Maschinen”, schreibt Georges Canguilhem in einem Essay von 1948, “können nur rationale ‘Normen’ von Kohärenz und Vorhersehbarkeit behaupten”. Im Gegensatz dazu sind Lebewesen zur Improvisation fähig; in seinem kreativen Drang zu überleben “toleriert das Leben Monstrositäten”. Unter Betonung der großen Plastizität des Nervensystems stellt Canguilhem fest, dass ein Kind, das einen Schlaganfall erleidet, der die Hälfte seines Gehirns zerstört, nicht an Aphasie leiden wird (wie es bei vielen Hirnverletzungen im späteren Leben der Fall ist), weil das Gehirn die Sprachfähigkeit auf andere Regionen umlenkt, um den erlittenen Schaden zu kompensieren. Organismen sind als Einheiten gerade deshalb stabil, weil ihre Organisation nicht in einer starren Struktur fixiert ist; sie sind offen und daher so ausgestattet, dass sie in manchen Fällen sogar einen Funktionsverlust überwinden können. Wie Canguilhem jedoch erklärte, “gibt es keine maschinellen Pathologien”[36]. Und natürlich offenbart für Canguilhem der pathologische Zustand der Krankheit die Macht des Lebewesens, in Krisensituationen neue Normen zu schaffen[37]. Michael Polanyi knüpfte an diese Beobachtung an, als er

schrieb: “Die Organe des Körpers sind komplexer und variabler als die Teile einer Maschine, und auch ihre Funktionen sind weniger klar spezifizierbar. Er stellte fest, dass eine Maschine nur “Regeln der Richtigkeit” kennt und sich niemals mit Fehlern oder “Ausfällen” abfinden kann – im Gegensatz zu lebenden Körpern, die zu radikalen Veränderungen fähig sind[38]. Oder, wie der Systemtheoretiker Ludwig von Bertalanffy es ausdrückte: Im Gegensatz zu einem lebendigen und offenen System wäre ein “mechanisierter Organismus unfähig, sich nach Störungen selbst zu regulieren”[39].

Mit dem Aufstieg der Kybernetiker kam jedoch der Glaube an eine völlig neue Art von Maschinen auf, eine flexible und anpassungsfähige Technologie, die die vitale Improvisation des Organismus nachahmen würde. Doch kybernetische Maschinen (z. B. selbstgesteuerte Raketen und andere Servomechanismen) wurden scheinbar immer noch von der Logik der Automatik beherrscht, trotz ihrer Fähigkeit, ihr Verhalten durch negative Rückkopplungsschaltungen zu korrigieren. Die adaptiven Reaktionen auf Umweltveränderungen wurden vollständig von dem programmierten System bestimmt. Das kybernetische Wesen hatte spezifische eingebaute “Ziele” und eine feste Organisation von Sensoren, Prozessoren und Motoren, um es zu diesen Zielen zu führen. Wie ein französischer Neurologe 1955 sagte, “erschien die Starrheit der in Maschinen oder Robotern entfalteten elektrischen Mechanismen von vornherein sehr weit entfernt von der Variabilität, der Plastizität derjenigen des Nervensystems”[40]. Konnte eine kybernetische Maschine echte Plastizität aufweisen? Konnte eine solche Maschine echte Pathologien in dem Sinne aufweisen, wie Goldstein und Canguilhem diesen Begriff definierten?[41].

Der Kybernetiker W. Ross Ashby hat die Möglichkeit einer solchen pathologischen Maschine untersucht, die dann zu einem wirklich neuen und unerwarteten Verhalten fähig wäre. In einem Fragment eines Notizbuchs aus dem Jahr 1943

finden wir ihn bei der Lektüre von William James. Ashby interessierte sich besonders für die Passagen bei James, in denen die starre “Maschine” mit der organischen und offenen Struktur des Nervensystems verglichen wird. Wie wir gesehen haben, wies dieses System für James paradoxerweise sowohl eine strukturelle Fixiertheit als auch eine adaptive, offene Plastizität auf[42]. Die Notizbücher zeigen auch, dass Ashby Charles Sherringtons bahnbrechende Arbeit über die neuronale Integration las, die 1906 veröffentlicht wurde[43]. Ashby konzentrierte sich auf die Passagen, in denen die dynamische Natur des Nervensystems beschrieben wurde. Wie James wies Sherrington auf den großen Vorteil der Fragilität des Nervensystems hin: Aufgrund seiner Störanfälligkeit reagiert der Organismus sehr empfindlich auf Veränderungen in seiner Umwelt[44]. Inspiriert von diesen Ideen machte sich Ashby daran, eine Maschine zu konstruieren, die sich wie ein offenes System verhalten könnte, d. h. wie eine determinierte Struktur, die dennoch in der Lage ist, sich neu zu organisieren.

Ashby räumte 1941 ein, dass künstliche Maschinen, die ihre Organisation ändern können, “selten”[45] sind. Er suchte nach der fehlenden Eigenschaft der Maschine, diesem “unberechenbaren Element” (in den Worten von James), das in der Lage ist, neue Handlungen hervorzubringen. Wenn Maschinen, wie Polanyi argumentiert hatte, durch ihre Rechenleistung definiert würden, wäre Ashbys Projekt unmöglich. Doch schließlich fand er einen neuen Weg, um über Maschinen und Plastizität nachzudenken. Im Wesentlichen erkannte er, dass eine bestimmte Maschine, um radikal offen für neue Formen der Organisation zu sein, in der Lage sein musste, in gewisser Weise eine völlig andere Maschine zu werden. Das Versagen oder der Ausfall, der in jeder konkreten Maschine bereits unvermeidlich war, erwies sich als Schlüsselidee. Für Ashby ist “ein Ausfall eine Änderung der Organisation”[46]. Eine kybernetische Maschine, die ständig versucht, ein Gleichgewicht zu finden, kann manchmal in eine Situation geraten, in der die Aufrechterhaltung der Homöostase unmöglich wird. Wenn dieses homöostatische

Wesen jedoch so konstruiert wäre, dass es in solchen Momenten auf die eine oder andere Weise zerbrechen kann, wenn das Gleichgewicht nicht erreicht wird, dann könnte man (zumindest theoretisch) sagen, dass diese Maschine eine neue Organisation und damit neue mögliche Wege zum Gleichgewicht erlangt. Wie Ashby schreibt, "ändert sich nach einem Ausfall die Organisation und damit auch die Gleichgewichte. Dadurch erhält die Maschine neue Chancen, ein neues Gleichgewicht zu erreichen, oder, wenn dies nicht der Fall ist, erneut zu zerbrechen."[47] Die Maschine ist nicht mehr in der Lage, sich selbst zu stabilisieren. In der Sprache der damaligen Neurophysiologie könnten wir sagen, dass das interne Versagen der Maschine eine Form des Schocks für das System war. Ashby ebnete also den Weg für eine neue Form der kybernetischen Plastizität, die sich aus der grundlegenden Schwäche jeder Maschine ergab, nämlich der Unvermeidbarkeit des Versagens.

Diese Ideen über die physikalischen Formen des adaptiven Verhaltens veranlassten Ashby, das menschliche Gehirn neu zu überdenken. Für ihn war das Gehirn nicht nur dynamisch und flexibel in seiner Organisation; es war auch eine besondere Art von "Maschine", die aus einzelnen Elementen (Neuronen) bestand, die ihrem Wesen nach unzuverlässig waren, in dem Sinne, dass sie bei bestimmten Schwellenwerten ihre Funktion einstellten, bis sie sich wieder erholen und erneut beginnen konnten, sich mit anderen Neuronen zu verbinden. Aufgrund ihrer Latenzzeit waren die einzelnen Neuronen sozusagen ständig dabei, im Gesamtsystem aufzutauchen und wieder zu verschwinden. Die Herausforderung der Kybernetik bestand darin, diese Form des produktiven Scheiterns innerhalb künstlich konstruierter intelligenter Maschinen zu modellieren[48].

Die Verbindung zwischen der Hirnforschung und der Kybernetik bietet eine neue Möglichkeit, die Ursprünge der KI zu kontextualisieren. Turing eröffnete das Feld zumindest auf konzeptioneller Ebene, als er 1950 Computing Machinery and

Intelligence veröffentlichte und der Öffentlichkeit das berühmte Imitationsspiel vorstellte. Turings Projekt, wie es allgemein verstanden wird, bestand darin, darüber nachzudenken, wie ein digitaler Computer erfolgreich die menschliche Intelligenz simulieren könnte, indem er das Denken durch algorithmische Prozesse modelliert. Doch 1946, als Turing gerade dabei war, Großbritanniens ersten digitalen Computer mit gespeichertem Programm, die Automatic Computing Engine, zu entwickeln, schrieb er an Ashby. In diesem Brief gab Turing zu, dass er "mehr an der Möglichkeit interessiert war, Modelle der Gehirntätigkeit zu erzeugen", als an den praktischen Anwendungen der neuen, von seiner Arbeit inspirierten Digitalcomputer[49]. Sein Interesse an Modellen des Nervensystems und des Gehirns deutet in der Tat darauf hin, dass er sich von der strikten Vorstellung des Automatismus der Maschine, die er in seiner formalmathematischen Arbeit über die Berechenbarkeit eingeführt hatte, abwandte und sich für die dynamische Organisation und Reorganisation organischer Systeme interessierte.

1936 hatte sich Turing den Universalcomputer theoretisch als eine vollkommen automatisierte, autonome Maschine vorgestellt. Doch 1948, nachdem er während und nach dem Zweiten Weltkrieg viel Erfahrung mit Computern gesammelt hatte, dachte er zunehmend über die Beziehung zwischen dem Gehirn, der menschlichen Intelligenz und diesen neuen Computern nach. Turing stellte fest, dass per Definition jede Maschine "vollständig beschrieben" werden kann, indem man alle ihre möglichen Konfigurationen aufzählt: "Gemäß unseren Konventionen wird die 'Maschine' vollständig durch die Beziehung zwischen ihren möglichen Konfigurationen zu aufeinanderfolgenden Zeitpunkten beschrieben. Dies ist eine Abstraktion, die sich im Laufe der Zeit nicht ändern kann". Dennoch stellte er sich vor, dass, wenn etwas in diese Maschine eingreifen würde, diese im Wesentlichen verändert würde; sie würde dann zu einer neuen Maschine werden. Turing wies dann suggestiv darauf hin, dass auch Schüler sehr wohl als "Maschinen"

betrachtet werden könnten, die durch massive Eingriffe, insbesondere durch Unterricht und Kommunikation sowie durch Formen sensorischer Reize, ständig verändert werden[50].

Turing merkte an, “dass es völlig unfair wäre, von einer Maschine, die direkt aus der Fabrik kommt, zu erwarten, dass sie mit einem Universitätsabsolventen unter gleichen Bedingungen kämpft. Letzterer hat zwanzig Jahre oder länger mit Menschen zu tun gehabt. Dieser Kontakt hat sein Verhalten über diesen Zeitraum hinweg verändert. Seine Professoren haben sich bemüht, ihn zu verändern”. Die “Routinen” des Denkens und Handelns eines Schülers wurden “dem ursprünglichen Modell seines Gehirns übergestülpt”. Der Mensch wurde im Laufe dieses “Programmierprozesses” nicht zu einem Automaten. Im Gegenteil: Die offene Maschine, die solchen Veränderungen durch Einmischung unterliegt, ist diejenige, die zu echter Kreativität fähig wird. Mit neuen Routinen kann der Maschinenmensch laut Turing “neue Kombinationen dieser Routinen ausprobieren, leichte Variationen anbringen und sie auf neue Weise anwenden”[51]. Das Fehlen automatischer Routinen ist für echte Intelligenz entscheidend: “Wenn man von einer Maschine erwartet, dass sie unfehlbar ist, kann sie nicht intelligent sein. Der Maschine muss der Kontakt mit Menschen erlaubt sein, damit sie sich an deren Normen anpassen kann”. Intelligenz”, schreibt Turing, “besteht darin, von dem völlig disziplinierten Verhalten abzuweichen, das die Berechnung mit sich bringt”[52]. Die Intelligenz eines Computers (oder eines Menschen) wird also nicht an seiner Rechenleistung gemessen, sondern an seiner radikalen Offenheit für Interferenzen – von außen natürlich, aber bis zu einem gewissen Grad auch von innen.

Turings Hypothese war, dass das entstehende menschliche Gehirn als eine unorganisierte Maschine zu betrachten sei, die durch geeignetes “Interferenztraining” eine Organisation erlangt: “All dies deutet darauf hin, dass

die Großhirnrinde des Säuglings eine unorganisierte Maschine ist, die durch geeignetes Interferenztraining organisiert werden kann [...]. Ein menschlicher Kortex wäre praktisch nutzlos, wenn man nicht versuchen würde, ihn zu organisieren. Wenn also ein Wolf durch Mutation einen menschlichen Kortex erwirbt, gibt es wenig Grund zu der Annahme, dass er daraus einen selektiven Vorteil ziehen wird."[53] Die Frage, ob ein Wolf einen menschlichen Kortex erwirbt oder nicht, lässt sich nicht beantworten. Ein isoliertes und einsames menschliches Gehirn würde keine Fortschritte machen, denn es braucht ein soziales Umfeld und die Interaktion mit anderen Menschen, um zu lernen und zu kreieren. Turings Projekt für die KI war nicht, wie allgemein angenommen, ein Projekt, das darauf abzielte, die Operationen der Kognition zu sezieren und sie in programmierbare Routinen zu übersetzen. Vielmehr ging es darum, ein offenes und flexibles Gehirn zu modellieren, das sich in dem Maße verändert, wie Natur und Kultur auf es einwirken. “Unsere Hoffnung”, schrieb er, “ist, dass es im Gehirn eines Kindes so wenige Mechanismen gibt, dass etwas Ähnliches leicht programmiert werden kann."[54] Das ist die Hoffnung, die wir haben. Für ihn ist die menschliche Intelligenz, wie auch die Computerintelligenz, ihrem Wesen nach bereits künstlich, d. h. von außen auf die plastische Architektur des Gehirns/Computers aufgeprägt.

Turings amerikanischer Gegenpart, John von Neumann, war ebenfalls fasziniert von dem robusten und flexiblen Verhalten organischer Nervensysteme und des Gehirns. Von Neumann, der von 1945 bis 1951 zur Überwachung der Entwicklung eines elektronischen Computers am Institute for Advanced Study in Princeton ernannt wurde, ließ sich von der Neurophysiologie inspirieren. Die Tatsache, dass das Gehirn und das Nervensystem eine erstaunliche Robustheit an den Tag legten, stand seiner Meinung nach in starkem Kontrast zu der immensen Anfälligkeit der neuen Hochgeschwindigkeitscomputer, die damals aus unzuverlässigen mechanischen Relais, Telefonschaltern oder Vakuumröhren

konstruiert wurden. In seinen frühen Überlegungen zum Computer achtete von Neumann darauf, die Aufmerksamkeit auf die wesentlichen Unterschiede zwischen digitalen Maschinen und dem Nervensystem zu lenken[55]. Dennoch fühlte er sich selbst vom Nervensystem als Mittel angezogen, um den Computer als etwas mehr als ein sequentielles Rechengerät zu denken. Eine der wichtigsten Eigenschaften des natürlichen Kommunikations- und Kontrollsystems war seiner Meinung nach seine inhärente Flexibilität und damit seine Zuverlässigkeit: “Es ist nie sehr einfach, etwas im Gehirn zu lokalisieren, weil das Gehirn eine enorme Fähigkeit zur Reorganisation hat. Selbst wenn man eine Funktion in einem bestimmten Teil des Gehirns lokalisiert hat, kann man, wenn man diesen Teil entfernt, feststellen, dass das Gehirn sich reorganisiert hat, seine Verantwortlichkeiten neu zugewiesen hat und die Funktion wieder ausgeführt wird.”[56] In den letzten Jahren hat Neumann immer wieder versucht, die Funktion zu reorganisieren.

1948 hielt von Neumann seinen berühmten Vortrag The General and Logical Theory of Automata (Die allgemeine und logische Theorie der Automaten) auf dem Hixon-Symposium, das am Cal Tech stattfand. Diese interdisziplinäre Zusammenkunft zum Thema Gehirn ist einer der wichtigsten Gründungsmomente der Kybernetik. Einige der einflussreichsten Forscher der Psychologie waren dort anwesend, darunter holistische Denker wie Köhler und Lashley. Tatsächlich schien von Neumann in seinem eigenen Vortrag für diese Tradition empfänglich zu sein. Er war an der Entwicklung künstlicher Computermaschinen interessiert, die die Robustheit natürlicher und lebender Systeme nachahmten, und er sprach genau die Art von Verhalten des Nervensystems an, die von Neurophysiologen untersucht wurde. Wie er erklärte, ist der Organismus immer mit unvorhersehbaren, aber unvermeidbaren Fehlern konfrontiert. Wie hält sich insbesondere das Nervensystem trotz des Drucks durch pathologische Umstände? In der Diskussion von von Neumanns Artikel wurde angemerkt, dass

wir “nicht nur in der Lage sein müssen, die normale Funktion des Nervensystems wiederzugeben, sondern auch seine relative Stabilität in allen möglichen abnormalen Situationen”[57].

In von Neumanns Arbeiten ist der Fehler mit Plastizität verbunden. Wenn ein komplexer Computer flexibel und stabil gegenüber Fehlern und Misserfolgen sein soll, ist es wichtig, dass “der Fehler [...] nicht als ein fremder, fehlgeleiteter oder fehlgeleiteter Unfall, sondern als ein wesentlicher Teil des betrachteten Prozesses betrachtet wird”[58]. Auf die eine oder andere Weise musste der Computer die Fähigkeit haben, “automatisch” zu erkennen, wo Fehler auftraten, und gleichzeitig das gesamte System neu zu organisieren, um diese Fehler zu berücksichtigen. Es scheint offensichtlich, dass von Neumann direkt auf Schlüsselbegriffe aus der theoretischen Biologie anspielte, die Lashley als “Äquipotentialität” bezeichnete und Bertalanffy unter “Äquifinalität” verstand. Von Neumanns Arbeiten befassten sich mit der Frage, wie sich natürliche Organismen neu organisieren können, wenn sie mit etwas konfrontiert werden, das er als “Notstandsbedingungen”[59] bezeichnete.

Von Anfang an war der Digitalcomputer eine Maschine auf der Suche nach ihrer Definition. Mit anderen Worten: Der Computer musste das enthalten, was der Technikphilosoph Gilbert Simondon “einen Spielraum der Unbestimmtheit” nannte. Für Simondon kann die perfekte Maschine niemals vollkommen automatisch sein. Wenn sie offen für Informationen und damit zu echter Kommunikation fähig sein soll, dann muss die Maschine, wie der Organismus, in gewissem Maße plastisch sein: “Die wahre Perfektion der Maschinen [...] entspricht nicht einer Zunahme der Automatisierung, sondern im Gegenteil der Tatsache, dass die Funktionsweise einer Maschine eine gewisse Marge der Unbestimmtheit birgt”. Zwar ist der Mensch, wie Descartes schon vor Jahrhunderten erkannte, eine Art Automat, der durch Erziehung, Bildung, Sprache

und Kultur bestimmt wird, doch handelt es sich dabei um einen Automaten besonderer Art. Wie Simondon sagt, “ist der Mensch ein ziemlich gefährlicher Automat, der immer das Risiko eingeht, neue Strukturen zu erfinden und sich selbst zu geben”[60]. Die Herausforderung bei der Schaffung eines jeden intelligenten Automaten besteht darin, Unbestimmtheit, Störung, Scheitern und Fehler zu modellieren. Im Zeitalter der Automatisierung, in dem die Kognition selbst als automatischer Vorgang betrachtet wird, ist es doppelt wichtig zu zeigen, dass der Automatismus des denkenden Wesens auf einer grundlegenderen Plastizität beruht. Lewis Mumford erläuterte dies 1964: “Lassen Sie mich die Vorstellung in Frage stellen, dass die Automatisierung gewissermaßen ein Endziel ist. Hätte sich der menschliche Organismus nach diesem Prinzip entwickelt, hätten Reflexe und das autonome Nervensystem alle Gehirnfunktionen absorbiert und der Mensch wäre ohne einen einzigen Gedanken in seinem Kopf geblieben. Organische Systeme [haben] einen Handlungsspielraum, die Freiheit, Fehler zu machen und zu korrigieren, wenig begangene Wege zu erkunden, unvorhersehbare Zufälle zu integrieren, das Unerwartete zu antizipieren, das Unmögliche zu planen.”[61] Der Mensch ist in der Lage, das Unmögliche zu planen und zu tun, was er will.

Im Zeitalter der Automatisierung, in dem die in den digitalen Technologien verkörperte Automatik die konzeptuellen Grundlagen der Kognitionswissenschaften strukturiert, ist es wieder an der Zeit, die menschliche Natur mit dem, was nicht automatisch ist, neu zu artikulieren. Unsere digitalen Gehirne – Gehirne, die von Computern geformt und simuliert und zunehmend durch wiederholte Interaktionen mit unseren digitalen Prothesen geformt werden – werden ihre wahre Plastizität erst dann offenbaren, wenn sie die Macht wiederentdecken, ihre eigene Automatizität zu unterbrechen.

David Bates

Finden Sie den Originaltext unter https://escholarship.org/content/qt00w2666h/qt00w2666h.pdf.

[1]Die ersten Abschnitte dieses Essays basieren auf meinem Artikel “Thinking automaticity on the threshold of digital”, in Digital Studies: Organologie des savoirs et technologies of knowledge, ed. Bernard Stiegler (Paris: Editions FYP, 2014).

[2]Morana Ala č , Handling Digital Brains: A Laboratory Study of Multimodal Semiotic Interaction in the Age of Computers (Cambridge, MA: MIT Press, 2011).

[3]Thomas P. Trappenberg, Fundamentals of Computational Neuroscience (Oxford: Oxford University Press, 2010).

[4]Zur verflochtenen Geschichte der Kognitionswissenschaften und der künstlichen Intelligenz siehe Paul Edwards, The Closed World: Computers and the Politics of Discourse in Cold War America (Cambridge, MA: MIT Press, 1997); Jean-Pierre Dupuy, On the Origins of Cognitive Science: La mécanisation de l’esprit, trad. M. B. DeBevoise (Cambridge, MA: MIT Press, 2009).

[5]Human Brain Project, “Overview”,” https://www.humanbrainproject.eu/discover/the-project/overview (accessed February 12, 2015).

[6]U.S. Department of Health and Human Services, National Institutes of Health, “Brain Research through Advancing Innovative Neurotechnologies (BRAIN),” braininitiative.nih.gov (accessed February 12, 2015).

[7]Catherine Malabou, The Ontology of the Accident: An Essay on Destructive Plasticity, trans. Caroline Shread (London: Polity, 2012).

[8]Rajagopal Ananthanarayanan, Steven K. Esser, Horst D. Simon, and Dharmendra S. Modha, “The Cat Is Out of the Bag: Cortical Simulations with 10 9 Neurons, 10 13 Synapses,” Proceedings of the Conference on High Performance Computing Networking, Storage and Analysis, article 63, http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6375547 (accessed June 2, 2015).

[9]Paul Merolla et al., “A Digital Neurosynaptic Core Using Embedded Crossbar Memory with 45pJ per Spike in 45nm,” IEEE Custom Integrated Circuits Conference, September 2011. http://ieeexplore.ieee.org/stamp/ stamp.jsp?tp=&arnumber=6055294 (accessed June 2, 2015).

[10]Mrigank Sharad et al., “Proposal for Neuromorphic Hardware Using Spin Devices,” Cornell University Library, July 18, 2012, http://arxiv.org/abs/1206.3227v4 (accessed June 2, 2015).

[11]Dharmendra S. Modha et al., “Cognitive Computing,” Communications of the ACM 54 (2011): 65.

[12]B. G. Farley and W. A. Clark, “Simulation of Self-Organizing Systems by Digital Computer,” Transaction of the IRE (1954): 7.

aNd.: das ein bestimmtes Verhalten hat, eine endliche Anzahl von Möglichkeiten, das sich nicht unendlich weiterentwickelt.

[13]René Descartes, Treatise on Man, in The Philosophical Writings of Descartes, trans. John Cottingham, Robert Stoothoff, and Dugald Murdoch, 3 Bde. (Cambridge: Cambridge University Press, 1985), 1: 104.

[14]René Descartes, Discourse on Method, in Philosophical Writings, 1: 119.

[15]David Bates, “Cartesian Robotics”, Representations 124 (Fall 2013): 43-68.

[16]C. S. Peirce, Collected Papers of Charles Sanders Peirce, 8 Bände (Cambridge, MA: Harvard University Press, 1931-58), 2: para. 59.

[17]C. S. Peirce, letter to A. Marquand, December 30, 1886, in C. Kloesel et al., eds, Writings of Charles S. Peirce: A Chronological Edition (Bloomington: Indiana University Press, 1993), 5: 421-22.

[18]William James, Principles of Psychology, 2 Bände (New York: Henry Holt, 1890), 1: 105.

[19]Ibid., 1: 139-40, 143.

[20]C. S. Peirce, Contributions to the Nation, 3 Bände (Lubbock: Texas Tech University Press, 1975-87), 1: 144.

[21]Constantin von Monakow, Die Lokalisierung im Großhirn und der Abbau der Funktion durch kortikale Herde (Wiesbaden: J. F. Bergmann, 1914), 30.

[22]Constantin von Monakow und R. Mourgue, Introduction biologique à l’étude de la neurologie et désintégration de la fonction (Paris: Félix Alcan, 1928), 29.

[23]Die folgende Diskussion über die Neurophysiologie greift die Argumentation in meinem Artikel auf: “Unity, Plasticity, Catastrophe: Order and Pathology in the Cybernetic Era,” in Catastrophes: History and Theory of an Operative Category, ed. Andreas Killen and Nitzen Lebovic (Berlin: De Gruyter, 2014), 32-54.

[24]Karl Lashley, Brain Mechanisms and Intelligence: A Quantitative Study of Injuries to the Brain (Chicago: University of Chicago Press, 1929), 176.

[25]Karl Lashley, “Persistent Problems in the Evolution of Mind”, QuarterlyReview of Biology 24 (1949): 33, 31.

Der sowjetische Psychologe und Neurologe Alexander Luria entwickelte eine ähnliche Theorie der neuronalen Plastizität. In seinen Studien an Patienten mit verschiedenen Hirnverletzungen beobachtete Luria den Verlust bestimmter kognitiver Fähigkeiten, stellte aber auch wiederholt fest, dass das Gehirn die erstaunliche Fähigkeit besitzt, sich neu zu organisieren, um den Verlust von Funktionen nach einem Schlaganfall oder Unfall zu kompensieren. Luria stellte fest, dass viele Studien "den hohen Grad an Plastizität beschädigter Funktionssysteme zeigen, der auf die dynamische Reorganisation und Anpassung an neue Umstände zurückzuführen ist und nicht auf die Regeneration und Wiederherstellung ihrer morphologischen Integrität". A. R. Luria, Restoration of Function after BrainInjury, trans. Basil Aigh (New York: Macmillan, 1963), 33. Siehe auch den Aufsatz von Laura Salisbury und Hannah Proctor in diesem Band.

[26]Kurt Goldstein, The Organism: A Holistic Approach to Biology Derived from Pathological Data in Man (1934; New York: Zone, 1995), 48-49. Ein ausgezeichneter Überblick über Goldsteins Arbeit im deutschen Kontext findet sich in Anne Harrington, Reenchanted Science: Holism in German Culture from Wilhelm II to Hitler (Princeton, NJ: Princeton University Press, 1996).

[27]Goldstein, The Organism, 52, 35.

[28]Ibid., 227, 388.

bNdt: Einsicht

[29]Wolfgang Köhler, Intelligenzprüfungen an Menschenaffen (Berlin: Springer, 1921).

[30]Wolfgang Köhler, “Some Gestalt Problems,” in A Sourcebook of Gestalt Psychology, ed. Willis D. Ellis (London: Kegan Paul, 1938), 55, 66.

[31] Karl Duncker, Zur Psychologie des produktiven Denkens (Berlin: Springer, 1935); Max Wertheimer, Productive Thinking (Ann Arbor: University of Michigan Press, 1945).

[32] Kurt Koffka, The Growth of the Mind: An Introduction to Child Psychology (New York: Harcourt Brace, 1925), 125.

[33] Edouard Claparède, “La genèse de l’hyopthèse: Étude expérimentale,” Archives de Psychologie 24 (1934): 5.

[34]Vladimir Jankélévitch, Henri Bergson (Paris: Presses Universitaires de France, 1989).

[35]Warren McCulloch und Walter Pitts, “A Logical Calculus of the Ideas Immanent in Nervous Activity,” Bulletin of Mathematical Biophysics 5 (1943): 115-33; Norbert Wiener, Cybernetics (Cambridge, MA: MIT Press, 1948); Warren McCulloch und John Pfeiffer, “Of Digital Computers Called Brains”, Scientific Monthly 69 (1949): 368-76.

[36]Georges Canguilhem, “Machine et organisme” (1948), in La connaissance de la vie, 2. Aufl. (Paris: Vrin, 1980), 118.

[37]Georges Canguilhem, On the Normal and the Pathological, trans. Carolyn R. Fawcett (New York: Zone, 1991).

[38]Michael Polanyi, Personal Knowledge: Towards a Post-Critical Philosophy (Chicago: University of Chicago Press, 1958), 359, 328.

[39]Ludwig von Bertalanffy, Das Biologische Weltbild (Bern: A. Franke, 1949), 29.

[40]Theophile Alajouanine, Ist der Mensch nur ein Roboter? Betrachtungen über die Bedeutung der Automatik in den Nervenfunktionen. Eröffnungsrede, gehalten auf der Sitzung des Kongresses der französischsprachigen Psychiater und Neurologen in Nizza am 5. September 1955 (Cahors: A. Coueslant, 1955), 11.

[41]Wir können feststellen, dass Bertalanffy die Kybernetik gerade wegen ihrer fehlgeleiteten Verwendung geschlossener Systeme zur Modellierung der grundsätzlich offenen Strukturen natürlicher Organismen kritisiert hatte. Ludwig von Bertalanffy, General System Theory: Foundations, Development, Applications (New York: George Braziller, 1968).

[42]W. Ross Ashby, Journal, 1523-24, W. Ross Ashby Digital Archive, rossashby.info/index.html (accessed July 24, 2013).

[43]Charles S. Sherrington, The Integrative Action of the Nervous System (New Haven, CT: Yale University Press, 1906).

[44]Ashby, Journal, 1906-9.

[45]Ibid., 1054.

[46]W. Ross Ashby, "The Nervous System as Physical Machine: With Special Reference to the Origin of Adaptive Behaviour," Mind 56 (1947): 50. Zu Ashbys theoretischer Arbeit und technischen Projekten in der Heimostase siehe Andrew Pickering, The Cybernetic Brain: Sketches of Another Future (Chicago: University of Chicago Press, 2009), insbesondere Kapitel 4.

[47]Ashby, "The Nervous System as Physical Machine", 55.

[48]Ibid., 57-58.

[49]Alan Turing, Brief an W. Ross Ashby, ca. November 19, 1946, zitiert in Jack Copeland, Einleitung zu Alan Turing, “Lecture on the Automatic Computing Engine (1947),” in Alan Turing, The Essential Turing: Seminal Writings in Computing, Logic, Philosophy, Artificial Intelligence, and Artificial Life plus the Secrets of Enigma, ed. B. Jack Copeland (Oxford: Oxford University Press, 2004), 374.

[50]Alan Turing, “Intelligent Machinery” (1948), in Turing, The Essential Turing, 419.

[51]Ibid., 421.

[52]Alan Turing, “Computing Machinery and Intelligence”, Mind 59 (1950): 459.

[53]Turing, “Intelligent Machinery”, 424.

[54]Turing, “Computing Machinery”, 456.

[55]John von Neumann, The Computer and the Brain (New Haven, CT: Yale University Press, 1958).

[56]John von Neumann, Theory of Self-Reproducing Automata, ed. Arthur W. Burks (Urbana: University of Illinois Press, 1966), 49.

[57]Ibid., 323.

[58]Ibid., 329. Ich habe von Neumanns Theorie des Fehlers in “Creating Insight: Gestalt Theory and the Early Computer,” in Genesis Redux: Essays in the History and Philosophy of Artificial Life, ed. Jessica Riskin (Chicago: University of Chicago Press, 2006), 237-59. Siehe auch Giora Hon, “Living Extremely Flat: The Life of an Automaton. John von Neuman’s Conception of Error of (In)animate Systems,” in

Going Amiss in Experimental Research, ed. Giora Hon, Jutta Schickore, and Friedrich Steinle, Boston Studies in the Philosophy of Science, no. 267 (N.p.: Springer, 2009), 55-71.

[59]Von Neumann, Theory of Self-Reproducing Automata, 71, 73.

[60]Gilbert Simondon, Du mode d’existence des objets techniques, rev. ed. (Paris: Aubier, 1989), 11-12.

[61]Lewis Mumford, “The Automation of Knowledge: Are We Becoming Robots?,” Audio-Visual Communication Review 12 (1964): 275.

Suche nach

META

All Topics

Authors

Datenschutzerklärung

Impressum

MORE MEDIA

    

ARCHIVE

Monat auswählen